CHAPTER – TWO

BACKGROUND AND LITERATURE SURVEY

2.1 Overview

This chapter presents the prerequisite and the conceptual that ties together all the key ideas. By "conceptual structure", it is meant that the organized way of linking individual ideas together so that their relative importance and interrelations are clear and obvious.

2.2 Literature Review

Predictive modelling uses statistics to predict outcomes. Most often the event one wants to predict is in the future, but predictive modelling can be applied to any type of unknown event, regardless of when it occurred. For example, predictive models are often used to detect crimes and identify suspects, after the crime has taken place.

In many cases, the model is chosen on the basis of detection theory to try to guess the probability of an outcome given a set amount of input data, for example given an email, to determine how likely that it is spam.

Models can use one or more classifiers in trying to determine the probability of a set of data belonging to another set, say spam or 'ham'.

Depending on definitional boundaries, predictive modelling is synonymous with, or largely overlapping with, the field of machine learning, as it is more commonly referred to in academic or research and development contexts. When deployed commercially, predictive modelling is often referred to as predictive analytics.

Some of the applications where predictive analysis is used are as follows:

- Uplift modelling
- Archaeology
- Customer relationship management
- Auto insurance
- Health care

2.3 Software requirement Specifications

A software requirements specification (SRS) is a document that captures complete description about how the system is expected to perform.

2.3.1 Functional Requirements

- Data refinement should be done properly while reading data from the dataset
- Relevant variables should be selected for model creation
- Model should be robust in nature
- Accuracy in the prediction should be at least 80%

2.3.2 Non-Functional Requirements

2.3.2.1 Execution Qualities

- Secure
- Easily Usable
- Robustness

2.3.2.2 Evolution Qualities

- Easily Maintainable
- Testable
- Scalable
- Accuracy of the results

2.4 Feasibility Study

A feasibility study aims to objectively and rationally uncover the strengths and weaknesses of an existing business or proposed venture, opportunities and threats present in the natural environment, the resources required to carry through, and ultimately the prospects for success.

2.4.1 Time Feasibility

Time feasibility study shows that time requires to create predictive models are directly proportional to the size of data set and the degree of the authenticity of the data set. Time required also depends on the hardware and software capabilities the system usd for the same purpose. From the past records, it has been observed that data sets with average size can be analysed and model generations are feasible in the real time.

2.4.2 Technical feasibility

Technical feasibility study inferred that predictive modelling where datasets are of average size can be done on personal computers while if datasets are large than use of cloud computing and Hadoop technology is to be employed in order to the predictive analysis.

2.4.3 Economic Feasibility

This project in particular is economically feasible while if data size tis to be humungous and algorithmic complexity of the program is large than predictive modelling can cost a large sum of money due to involvement of cloud technology.

2.4.4 Legal Feasibility

There is no legal issues with the project as the dataset is in the public domain and implementation of the modelling techniques is open source and in the public domain.

2.4.5 Operational Feasibility

It deals with the consideration about working of the system after installation. The proposed system would be beneficial to its users as their needs are fully satisfied. As this project satisfies all the requirements of the users, it is operationally feasible. All the operational aspects are considered carefully here. Only by spending tie to evaluate feasibility, we will be able to reduce the chances for extreme embracement at later stages of a project.
The benefits of proposed system are:

- Ability to handle significantly large amount of a data
- Fast and accurate information is possible
- Easy report generation

2.4.6 Social Feasibility

The loan predictor models prove to be a great asset for the banking sector and they are employing the use of his technology to a lager extent. But if the model is to be made available in the public domain then it has a great potential of exploitation of the business model of the banking sector and can pose a serious threat to the credibility of the banks and a result of that they can lose a large no of customers.

2.5 Innovativeness and Usefulness

Predictive modelling brilliantly uses the mathematical models to predict the outcome given a number of parameters as input.

Usefulness of a loan predictor models are as follows:

- Helps banks to predict whether the applicant's chances of success in securing the loan from the bank
- If the predicted chances of loan approval is very low then banks do not waste their time in the processing of the application and cancels the candidature for the same
- If the predicted chances of loan approval is very high then banks do not waste their time in the processing of the application and accelerates the process
- Saves bank a lots of money that goes into the documents processing